



**HAL**  
open science

## Enhancing fetal alcohol spectrum disorders diagnosis with a classifier based on the intracerebellar gradient of volumetric undersizing

Justine Fraize, Clara Fischer, Monique Elmaleh-Bergès, Eliot Kerdreux, Anita Beggiato, Alexandra Ntorkou, Edouard Duchesnay, Dhaif Bekha, Odile Boespflug-Tanguy, Richard Delorme, et al.

### ► To cite this version:

Justine Fraize, Clara Fischer, Monique Elmaleh-Bergès, Eliot Kerdreux, Anita Beggiato, et al.. Enhancing fetal alcohol spectrum disorders diagnosis with a classifier based on the intracerebellar gradient of volumetric undersizing. *Human Brain Mapping*, 2023, 44 (11), pp.4321-4336. 10.1002/hbm.26348 . inserm-04102112

**HAL Id: inserm-04102112**

**<https://inserm.hal.science/inserm-04102112>**

Submitted on 22 May 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.





Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

## RESEARCH ARTICLE

WILEY

# Enhancing fetal alcohol spectrum disorders diagnosis with a classifier based on the intracerebellar gradient of volumetric undersizing

Justine Fraize<sup>1,2</sup>  | Clara Fischer<sup>3</sup> | Monique Elmaleh-Bergès<sup>2,4</sup> |  
 Eliot Kerdreux<sup>1,2</sup> | Anita Beggato<sup>5</sup> | Alexandra Ntorkou<sup>4</sup> |  
 Edouard Duchesnay<sup>3</sup>  | Dhaif Bekha<sup>1,2</sup> | Odile Boespflug-Tanguy<sup>6</sup> |  
 Richard Delorme<sup>5</sup> | Lucie Hertz-Pannier<sup>1,2</sup> | David Germanaud<sup>1,2,7</sup>

<sup>1</sup>CEA Paris-Saclay, Joliot Institute, NeuroSpin, UNIACT, Centre d'études de Saclay, Gif-sur-Yvette, France

<sup>2</sup>Université Paris Cité, Inserm, U1141 NeuroDiderot, inDEV, Paris, France

<sup>3</sup>CEA Paris-Saclay, Joliot Institute, NeuroSpin, BAOBAB, Centre d'études de Saclay, Gif-sur-Yvette, France

<sup>4</sup>Department of Pediatric Radiology, Centre of Excellence InovAND, AP-HP, Robert-Debré Hospital, Paris, France

<sup>5</sup>Department of Child and Adolescent Psychiatry, Centre of Excellence InovAND, AP-HP, Robert-Debré Hospital, Paris, France

<sup>6</sup>Université Paris Cité, U1141 NeuroDiderot, NeuroDEV, Paris, France

<sup>7</sup>Department of Genetics, Centre of Excellence InovAND, AP-HP, Robert-Debré Hospital, Paris, France

## Correspondence

Justine Fraize and David Germanaud, CEA Paris-Saclay, Joliot Institute, NeuroSpin, UNIACT, Centre d'études de Saclay, Bâtiment 145, 91191, Gif-sur-Yvette, France.  
 Email: [justine.fraize@inserm.fr](mailto:justine.fraize@inserm.fr) and [david.germanaud@cea.fr](mailto:david.germanaud@cea.fr)

## Funding information

French National Agency for Research; French National Institute for Public Health research

## Abstract

In fetal alcohol spectrum disorders (FASD), brain growth deficiency is a hallmark of subjects both with fetal alcohol syndrome (FAS) and with non-syndromic FASD (NS-FASD, i.e., those without specific diagnostic features). However, although the cerebellum was suggested to be more severely undersized than the rest of the brain, it has not yet been given a specific place in the FASD diagnostic criteria where neuro-anatomical features still count for little if anything in diagnostic specificity. We applied a combination of cerebellar segmentation tools on a 1.5 T 3DT1 brain MRI dataset from a monocentric population of 89 FASD (52 FAS, 37 NS-FASD) and 126 typically developing controls (6–20 years old), providing 8 volumes: cerebellum, vermis and 3 lobes (anterior, posterior, inferior), plus total brain volume. After adjustment of confounders, the allometric scaling relationship between these cerebellar volumes ( $V_i$ ) and the total brain or cerebellum volume ( $V_t$ ) was fitted ( $V_i = bV_t^a$ ), and the effect of group (FAS, control) on allometric scaling was evaluated. We then estimated for each cerebellar volume in the FAS population the deviation from the typical scaling ( $\Delta$ DTS) learned in the controls. Lastly, we trained and tested two classifiers to

**Abbreviations:** AIC, Akaike information criterion; BIC, Bayesian information criterion;  $\Delta$ DTS, deviation from the typical scaling; FASD, fetal alcohol spectrum disorder; FAS, fetal alcohol syndrome; FDR, false discovery rate; MRI, magnetic resonance imaging; NDD, neurodevelopmental disorders; NS-FASD, non-syndromic FASD; PAE, prenatal alcohol exposure; ROC, receiver operating characteristic.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Human Brain Mapping* published by Wiley Periodicals LLC.

discriminate FAS from controls, one based on the total cerebellum  $\sqrt{\text{DTS}}$  only, the other based on all the cerebellar  $\sqrt{\text{DTS}}$ , comparing their performance both in the FAS and the NS-FASD group. Allometric scaling was significantly different between FAS and control group for all the cerebellar volumes ( $p < .001$ ). We confirmed the excess of total cerebellum volume deficit ( $\sqrt{\text{DTS}} = -10.6\%$ ) and revealed an antero-inferior-posterior gradient of volumetric undersizing in the hemispheres ( $-12.4\%$ ,  $1.1\%$ ,  $2.0\%$ , respectively) and the vermis ( $-16.7\%$ ,  $-9.2\%$ ,  $-8.6\%$ , respectively). The classifier based on the intracerebellar gradient of  $\sqrt{\text{DTS}}$  performed more efficiently than the one based on total cerebellum  $\sqrt{\text{DTS}}$  only (AUC = 92% vs. 82%,  $p = .001$ ). Setting a high probability threshold for >95% specificity of the classifiers, the gradient-based classifier identified 35% of the NS-FASD to have a FAS cerebellar phenotype, compared to 11% with the cerebellum-only classifier ( $p_{\text{FISHER}} = 0.027$ ). In a large series of FASD, this study details the volumetric undersizing within the cerebellum at the lobar and vermian level using allometric scaling, revealing an anterior-inferior-posterior gradient of vulnerability to prenatal alcohol exposure. It also strongly suggests that this intracerebellar gradient of volumetric undersizing may be a reliable neuroanatomical signature of FAS that could be used to improve the specificity of the diagnosis of NS-FASD.

#### KEYWORDS

cerebellum morphometry, diagnostic imaging, fetal alcohol spectrum disorders, fetal alcohol syndrome, microcephaly, prenatal alcohol exposure, scaling analysis

## 1 | INTRODUCTION

The clinical consequences of prenatal alcohol exposure (PAE) are grouped under the diagnostic continuum of fetal alcohol spectrum disorder (FASD), ranging from the specific form called fetal alcohol syndrome (FAS) to nonspecific forms called non-syndromic FASD (NS-FASD). FASD diagnosis relies on four criteria: growth deficiency, facial dysmorphism, anatomical (mainly brain growth deficiency) and functional neurological abnormalities, in conjunction with PAE. In NS-FASD, the physical features of FAS are absent or insufficient, and the probabilistic diagnosis is based on the combination of compatible neurodevelopmental disorders (NDD) and sufficient PAE. This distinction and the diagnostic criteria of these two conditions were formalized by Astley (2004) in the 4-digit diagnostic code, by Hoyme et al. for the Institute of Medicine (Hoyme et al., 2016) and in the Canadian guidelines by Cook et al. (2016). Yet there is still a need for additional markers particularly to ensure a high likelihood of causal link between observed NDD and PAE in the absence of FAS, and thus support the use of the whole diagnosis spectrum, as NS-FASD are still underdiagnosed.

The most obvious and widely reported consequence of PAE on the brain is growth deficiency (Archibald et al., 2001; Astley et al., 2009; Boronat et al., 2017; Nardelli, Lebel, Rasmussen, Andrew, & Beaulieu, 2011; Rajaprakash, Chakravarty, Lerch, & Rovet, 2014; Roebuck, Mattson, & Riley, 1998; Treit et al., 2014; Treit, Jeffery, Beaulieu, & Emery, 2020). It can result in microcephaly, a clinical feature

that may (Hoyme et al., 2016), or may not (Astley, 2004; Cook et al., 2016), be considered a distinct diagnostic criterion. Although the brain is one of the main targets of alcohol developmental toxicity, no other neuroanatomical anomaly is specified in the diagnostic guidelines, let alone used to improve diagnostic certainty. Yet, some focal anomalies have now been reported in an interestingly high number of subjects with FAS or FASD, particularly cerebellar anomalies, initially just described on clinical magnetic resonance imaging (MRI) (Astley et al., 2009; Boronat et al., 2017; Treit et al., 2020), and more recently specified by our team as at least vermian hypoplasia and/or anterior dysgeneses (Fraize et al., 2022).

MRI-based computational neuroanatomy has enabled the description of damage not identified by the radiologist but significant in quantitative group analysis, supporting the existence of and further specifying cerebellar involvement in FASD (Astley et al., 2009; Bookstein, Streissguth, Connor, & Sampson, 2006; Cardenas et al., 2014; O'Hare et al., 2005; Sowell et al., 1996). These first studies relied on small samples and on tools of increasing sophistication to characterize these abnormalities. Only in one recent study did Sullivan et al. (2020) implement a comprehensive automated segmentation of the cerebellum (yet without full individualization of anterior segments), in a large cohort of subjects with FASD, to characterize lower intracerebellar volume. After linear correction for sex, age and noteworthily brain size, they showed significantly lower volume of both the anterior part of the cerebellum and some posterior lobules (VIIIa and X lobules) in the FASD group compared to controls. Yet, despite a breakthrough in

cerebellar lesion mapping, no relevant clinical translation of these average group-level effects was proposed.

Indeed, any clinical application requires the analysis to result not only in group characterization but in individual assessment. At least, it should lead to the identification of subjects bearing neuroanatomical features that are both recurrent and distant enough from a normative reference to be considered as markers of FASD (such as the FAS clinical features); at best, it should provide a probability of being fetal alcohol-related. In a previous study, we identified several anomalies recurrent in FAS from a set of manual clinical-radiological measurements. We proposed to aggregate these multifocal neuroanatomical features into a composite phenotype to identify subjects with FASD with improved specificity (Fraize et al., 2022). With the same idea of multivariate analysis of neuroanatomical features for the discrimination of subjects with FASD, Little and Beaulieu (Little & Beaulieu, 2020), recently introduced a machine learning strategy considering a rather broad brain phenotype: using all supra-tentorial brain volumes in a classifier, they sought to identify brain regions with the greatest predictive contribution to FASD diagnosis but without FAS vs. non-FAS distinction.

In addition to sample size, type of anatomical segmentation or individualization of analysis, a final pitfall may have limited the success in identifying useful neuroanatomical markers of FASD: in most studies, the local size deficits were interpreted by taking global brain size into account with a linear or even sometimes a strictly proportional (ratio) model (Astley et al., 2009). Yet the relevance of the power law to describe the phenomena associated with brain size variations (scaling), especially age-independent size polymorphism, is well advocated (Bookstein et al., 2006; de Jong et al., 2017; Germanaud et al., 2012; Liu, Johnson, Long, Magnotta, & Paulsen, 2014). This scaling model allows for variation in shape and proportions with brain size (allometric scaling), including the strictly proportional model as a special case (isometric scaling with scaling exponent = 1). Even more, it has been shown that not using an allometry-sensitive model could be misleading in the case of pronounced size effects (Germanaud et al., 2014; Toro et al., 2009). In the general population, the cerebellum has been shown to scale with very negative allometry with brain size, that is, larger brains have relatively smaller cerebellums (de Jong et al., 2017; Warling et al., 2021). This nonlinear size effect should be considered in any comparison between FASD versus typically developing individuals, where brain size ranges only partially overlap, to interpret the volume variations of affected subjects according to the typical scaling of controls.

In this study, we searched for volumetric abnormalities in the cerebellum, vermis, and their lobar sub-parts, in FASD individuals compared with typically developing controls, taking into account the allometric cerebellar scaling, so as to exploit the detailed profile of cerebellar and intracerebellar anomalies to better characterize subjects with FASD. After proper anatomical segmentation, the scaling laws of the cerebellar volume (as a function of total brain volume) and each of its sub-volumes (as a function of the cerebellar volume) were estimated in the typically developing group, to be compared to the scaling laws estimated in the FASD group, but mostly to serve as

references to evaluate the deviation from typical scaling of these volumes in each subject with FASD. This normative-like approach aimed at both estimating a comprehensive profile of cerebellar volumetric undersizing, identifying the regions more or less affected by the global cerebral or cerebellar volume deficit, and providing individual profiles to be integrated in a logistic classifier of FAS versus typically developing controls. Lastly, we estimated the effectiveness of this FAS versus control classifier, and particularly whether it could help identify a significant subgroup of subjects with NS-FASD with a FAS-like cerebellar pattern.

## 2 | METHODS

### 2.1 | Participants

Eighty-nine consecutive subjects with FASD, aged 6 to 20 years, were retrospectively included from a clinical series of patients attending the NDD-dedicated child neurology consultation at Robert-Debré University Hospital (RD) between 2014 and 2020. FASD diagnostic was established using the two main guidelines (Astley, 2004; Hoyme et al., 2016) and a full differential diagnosis work-up was completed, including a systematic brain MRI. Individuals prenatally exposed to another embryo-fetotoxic agent were not included, nor were those who explicitly refused to participate in the study (see details in supplementary material Figure A.1). Subjects with FASD were split into two groups: the syndromic or FAS (including partial FAS) and the non-syndromic or NS-FASD. We evaluated the concordance (Cohen's kappa coefficient,  $\kappa$ ) of between-group assignment using one or the other guideline. This cohort and the diagnostic procedure have already been described in a previous study (Fraize et al., 2022). Clinical and radiological characteristics of the 52 subjects with FAS (58.4%) and the 37 subjects with NS-FASD (41.6%) are detailed in Table 1.

One hundred and twenty-six typically developing subjects, aged 6 to 20 years, with no report of PAE, developmental delay or family history of neurological or psychiatric condition (1st degree) were included for comparison (supplementary material Figure A.1 and Table A.1). A subgroup of 40 subjects was matched with the FASD group for the acquisition site (MRI scanner and sequence) as part of a research program on autism in the RD Psychiatry Department. Other typically developing subjects were part of previously published studies (Boueyre et al., 2018; Germanaud et al., 2014).

There were no significant differences in the control group compared to the FASD group for sex (50.8% vs. 58.4% of male respectively,  $p = .334$ ) and age at MRI (12.08 vs. 11.32 years of age respectively,  $p = .116$ ).

### 2.2 | Ethics statement

This study was conducted in accordance with the principles of the Declaration of Helsinki. Subjects' data were studied in accordance

**TABLE 1** Demographic, clinical, radiological data of subjects with FASD.

	FAS (n = 52)	NS-FASD (n = 37)	FASD groups comparison (p-value)
Sociodemographic assessment			
Sex: male, n (%)	27 (51.9)	25 (67.6)	.209
Age at MRI, mean in years (SD)	10.93 (3.57)	11.88 (3.55)	.219
Clinical assessment, n (%)			
(1) Prenatal alcohol exposure			
4.Confirmed, severe	21 (40.4)	16 (43.2)	.959
3.Confirmed, moderate or unquantified	26 (50.0)	19 (51.1)	1.000
2.Not documented	5 (9.6)	2 (5.4)	.748
1.No exposure	0 (0.0)	0 (0.0)	–
(2) FAS facial features			
4.Severe	31 (59.6)	2 (5.4)	<b>&lt;.001</b>
3.Moderate	21 (40.3)	1 (2.7)	<b>&lt;.001</b>
2.Mild	0 (0.0)	30 (81.1)	<b>&lt;.001</b>
1.None	0 (0.0)	4 (10.8)	.057
(3) Growth deficiency			
4.Significant	19 (36.5)	3 (8.1)	<b>.005</b>
3.Moderate	11 (21.2)	2 (5.4)	.077
2.Mild	9 (17.3)	9 (24.3)	.586
1.None	13 (25.0)	23 (62.2)	<b>.001</b>
Brain anatomy			
(4) Structural central nervous system damage	40 (76.9)	19 (51.4)	.290
Head circumference (smallest known)			
(4) $\leq -2$ SD: microcephaly	34 (65.4)	13 (35.1)	<b>.009</b>
Diagnostic agreement with revised Institute of Medicine Guidelines	50 (96.2)	33 (89.2)	
Cohen's Kappa coefficient	$\kappa = 0.86$		

Note: In bold, p-values <0.05.

Abbreviations: FAS, fetal alcohol syndrome; FASD, fetal alcohol spectrum disorder; NS-FASD, non-syndromic fetal alcohol spectrum disorder; SD, standard deviation.

with French regulation (MR-004, declaration of conformity n°2059980v0), following approval by the Paris-Saclay research ethics committee (CER-Paris-Saclay-2020-094). Controls' data were used within the framework of the ethical authorizations of the primary studies (Gene and autism, Inserm C07-33, 08-029 and 11-008).

## 2.3 | MRI acquisition

For both subjects with FASD and site-matched controls, MRI acquisitions were performed in the Department of Pediatric Radiology of RD Hospital at 1.5 T (Ingenia, Philips Healthcare, Amsterdam, the Netherlands) with a 3DT1 FFE-TFE sequence (1 mm isotropic; TR = 8.2 ms; TE = 3.8 ms; TI = 0.8 s; Flip = 8°; SENSE = 2). Another group of controls was acquired at the Frédéric Joliot Hospital (SHFJ, CEA-Saclay) at 1.5 T (Signa, GE Healthcare, Milwaukee, US) with a 3DT1 GE FSPGR sequence (1 × 1 × 1.2 mm; TR = 9.9 ms; TE = 2 ms; TI = 0.6 s; Flip = 10°) and a third one on a 3 T (Siemens

Trio, Siemens Healthineers, Oxford, UK) at NeuroSpin (NS, CEA-Saclay) with a 3DT1 Siemens MPRAGE sequence (1 mm isotropic; TR = 2.3 s; TE = 3 ms; TI = 0.9 s; Flip = 9°; GRAPPA 2). A visual quality check was systematically performed (Justine Fraize and David Germanaud) to exclude images of insufficient quality (detailed in supplementary material Figure A.1).

## 2.4 | MRI processing and cerebellar segmentation

We applied the automated segmentation of the cerebellum into lobules proposed by CERES (Romero et al., 2017), in volBrain (<http://volbrain.upv.es>, Manjón & Coupé, 2016). CERES is based on multi-atlas label fusion, known to minimize labelling errors due to inaccurate affine or non-linear registration. CERES has proven its superiority over other methods (Carass et al., 2018) but does not individualize the vermis. Thus, we also extracted the mask of the vermis from the single individual reference “Scalable Brain Atlas” (Bakker, Tiesinga, &

Kötter, 2015) and for each subject, we applied *SUIT* (Diedrichsen, Balsters, Flavell, Cussans, & Ramnani, 2009), a spatial normalization tool optimized for the cerebellum, to perform a mono-atlas segmentation based on our high-resolution atlas of the vermis. Then we merged the mask of the vermis with the *CERES* segmentation in each subject space. We extended the lobules' "cortical" labels in the deep cerebellum by applying a Voronoi front propagation algorithm available in the *Morphologist* software suite (*BrainVISA*) (<https://brainvisa.info>). We finally combined lobules to obtain sub-volumes, according to Schmahmann's anatomical description (Schmahmann et al., 1999) and merged the right and left sides, providing eight volumes and sub-volumes: the whole cerebellum, whole vermis and their subdivisions into anterior, posterior, and inferior lobes (hereinafter referred to as anterior, posterior and inferior hemisphere or vermis). All steps were followed by a visual quality check and manual corrections of obvious errors of segmentation (Figure 1). The total brain volume was obtained by *volBrain* (Manjón & Coupé, 2016).

## 2.5 | Statistical analysis

The statistics were computed using the R software (R Foundation for Statistical Computing, Vienna, Austria). Subject characteristics and volumes were analyzed using proportion test or independent samples

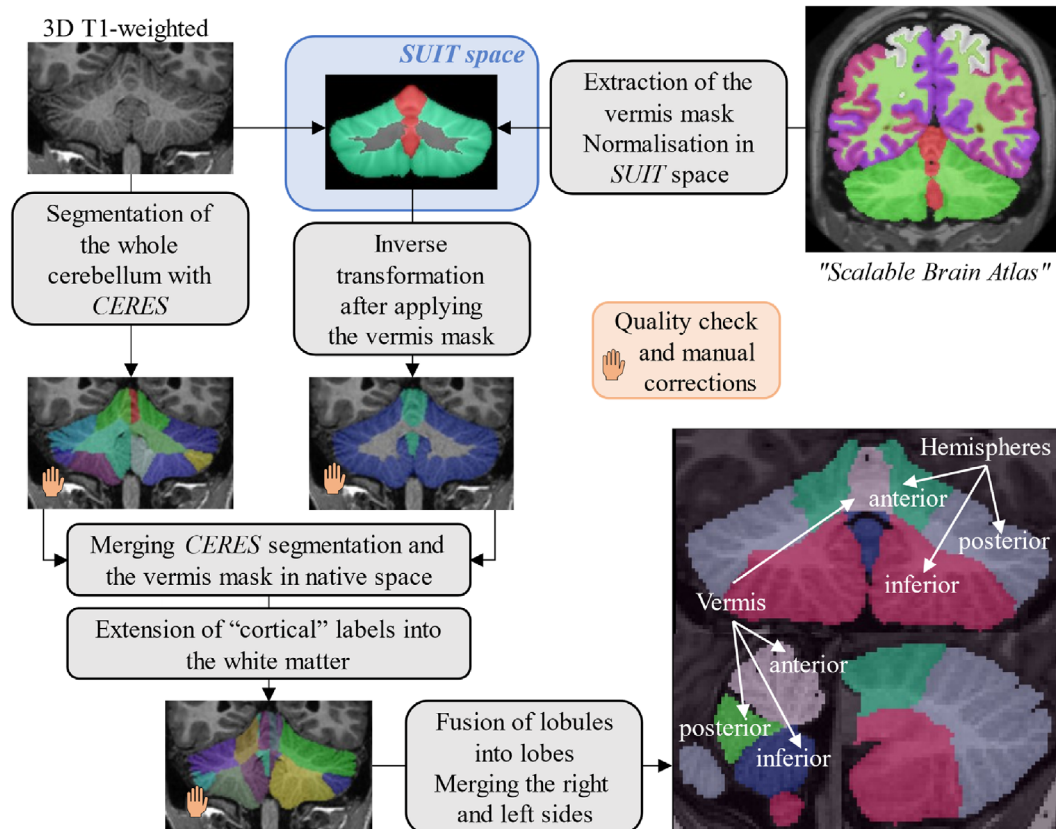
t-tests. For each analysis step, p-values were corrected to achieve a false discovery rate (FDR) of 5% (Benjamini & Hochberg, 1995).

### 2.5.1 | Adjusting for confounders

Raw volumes were corrected for confounders (site, sex, age) using ComBat (Fortin et al., 2018) based on an empirical Bayes method for correction of batch effects (Johnson, Li, & Rabinovic, 2007). We sequentially corrected for site (RD, SHFJ, NS) then for sex, integrating into the model correction sex, age (quadratic effect) and diagnostic (FASD, control) as covariates to be spared. The residual effect of age was no longer significant and was not corrected. In a robustness *post-hoc* analysis, we performed all analyses described below with the raw unadjusted data.

### 2.5.2 | Evaluating the effect of FAS on scaling

We performed scaling analyses modeling both the effect of brain size on the cerebellar volume, and at the cerebellar scale, the effect of cerebellum size (total cerebellum volume) on the cerebellar sub-volumes. We used a non-linear power law that accounts for allometric effects, i.e. the variation of proportion with size by comparison with a strictly



**FIGURE 1** Cerebellar segmentation pipeline.



proportional relationship called isometry (Bookstein et al., 2006; de Jong et al., 2017; Germanaud et al., 2012; Liu et al., 2014). To assess this scaling relationship between cerebellar volumes ( $V_i$ ) and total volume ( $V_t$  = total brain or cerebellum volume), we fitted  $V_i = bV_t^a$ , in the FAS and the control groups, after log-linearization ( $\log_{10} V_i = a \log_{10} V_t + \log_{10} b$ ). Scaling coefficients 'a' in the FAS ( $a_{FAS}$ ) and control ( $a_{Controls}$ ) groups were compared to 1, to determine whether the fit was different from isometry. We evaluated the effect of FAS diagnosis on allometric laws by comparing nested models with and without a FAS diagnosis covariate ( $\log_{10} V_i = a \log_{10} V_t + a_{FAS} \text{ FAS}$   $\log_{10} V_t + \log_{10} b + \text{FAS} \log_{10} b_{FAS}$ ) with a Wald test. We also estimated the determination coefficient  $R^2$  of the global model and its difference with the model with the FAS diagnosis covariate  $\delta R^2$ .

### 2.5.3 | Quantifying the degree of volumetric undersizing in FASD compared to control

Using the scaling model fitted in the typically developing group as a reference, we established the individual deviations from the predicted value for each subject (FAS, NS-FASD and controls). As a metric of this deviation, we proposed a "deviation from typical scaling" ( $\sqrt{DTS}$ ), which corresponds to the distance from the typical/reference scaling prediction, relative to the mean volume of the controls, as  $\sqrt{DTS} = (V_{observed} - V_{norm})/\sqrt{V_{control}}$ . For the sake of neuroanatomical clarity, we mapped the  $\sqrt{DTS}$  values with a color scale on the cerebellar segmentation of selected subjects.

To perform normative analysis,  $\sqrt{DTS}$  and confidence intervals were evaluated in the FAS group taking into account both the variance of the estimation of the reference model in controls and the variance of the FAS observations with two embedded bootstrap resamplings ( $n = 1000$ ). We established the distribution of the mean difference between the  $\sqrt{DTS}$  in the control ( $\approx 0$  expected) and in the FAS group in a total of  $10^6$  bootstrapped observations. The proportion of observations in which the null hypothesis (controls-FAS  $\neq 0$ ,  $\alpha = 5\%$ ) was rejected among the bootstrapped observations enabled calculation of the  $p$ -value. The differences between each sub-part of the cerebellum were investigated two-by-two by estimating if the distribution of their difference in  $\sqrt{DTS}$  included 0 or not.

### 2.5.4 | Classifying the subjects with FASD

Based on  $\sqrt{DTS}$ , we fitted logistic regression models to estimate the probability of belonging to the FAS group. We built two nested models, the first using only the  $\sqrt{DTS}$  of the cerebellum (cerebellum-only classifier) and the second adding the six cerebellar sub-volumes  $\sqrt{DTS}$  (gradient-based classifier). The benefit of adding the intracerebellar  $\sqrt{DTS}$  to the classifier was tested successively in several ways. First, the fitting quality of the logistic regressions was tested by using the Wald test and comparing both the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). Then the inherent performance of the classifiers was evaluated in the whole FAS

+ control population by computing the area under the receiver operating characteristic (ROC) curves ( $AUC_{global}$ ) and comparing them with the DeLong method (DeLong, DeLong, & Clarke-Pearson, 1988). As our aim was to obtain a classifier able to improve the diagnosis likelihood, we decided that the specificity ought to be high, thus we set the one-decimal threshold probability for a diagnostic prediction of FAS to achieve a specificity  $>95\%$  at least, in both classifiers used in the whole population. Second, to further assess the performance of each classifier, we performed a 5-fold cross-validation (Mosteller, 1968). Over the five validation folds, we estimated the mean of the performance criteria ( $AUC_{CV}$ ) and of both effective specificity and accuracy under the previously defined diagnosis threshold.

Finally, each high-specificity classifier (cerebellum-only and gradient-based classifiers) fitted in the whole population was applied to the FAS and NS-FASD populations to estimate and compare the proportion of subjects classified as FAS on neuroanatomical features only, that is to say for subjects with NS-FASD those showing a FAS-like cerebellar or intracerebellar profile of  $\sqrt{DTS}$ .

## 3 | RESULTS

### 3.1 | Mean volume comparison between groups

Group of subjects with FAS or NS-FASD showed lower raw (data not shown) and adjusted volumes (total brain, total cerebellum, and all cerebellar sub-parts) compared to the control group (Figure 2). The volumes in NS-FASD group were intermediate between controls and FAS groups.

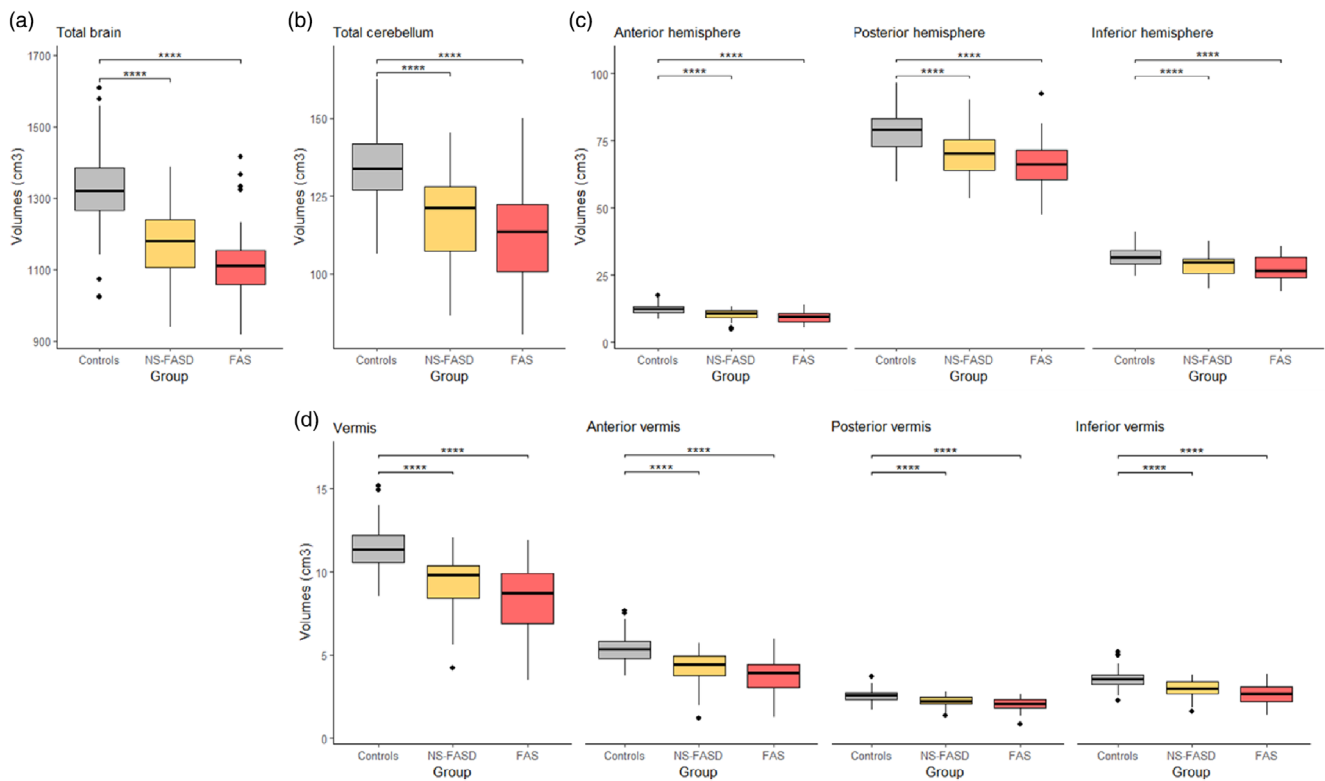
### 3.2 | Cerebellar scaling analysis: Accounting for the brain size effect

#### 3.2.1 | Reference scaling laws in controls

In controls, the global scaling of the cerebellum volume relative to the total brain volume showed the expected strong negative allometry ( $a_{control} = 0.40$ ,  $p < .001$ ), illustrated in Figure 3a. Relative to the cerebellum volume, the intracerebellar scaling of the sub-parts was not significantly different from isometry ( $a \approx 1$ ), except for the posterior vermis ( $a_{control} = 0.62$ ,  $p = .034$ ) (Figure 3).

#### 3.2.2 | Effect of FAS on cerebellar scaling laws

In the FAS group, the global scaling of the cerebellum volume relative to the total brain volume was isometric (i.e., proportional,  $a_{FAS} = 1.02$ ,  $p = .899$ ). The intracerebellar scaling analysis showed positive allometry for the anterior hemispheres ( $a_{FAS} = 1.38$ ;  $p = .013$ ) and its vermis ( $a_{FAS} = 2.05$ ;  $p < .001$ ), negative allometry for the posterior hemispheres ( $a_{FAS} = 0.86$ ;  $p = .001$ ) and isometry for the inferior



**FIGURE 2** Cerebral and cerebellar adjusted volumes for control, FAS and NS-FASD groups. (a) Total brain, (b) total cerebellum, (c) cerebellar hemispheres, (d) vermis and its sub-parts volumes. NS-FASD, non-syndromic fetal alcohol spectrum disorder. The difference between FASD and controls groups was significant for all the volumes (\*\*\*\* for  $p < .01$ ).

hemispheres ( $a_{FAS} = 1.08$ ;  $p = .279$ ). Some of the vermis sub-volumes exhibited a different scaling from that of their corresponding hemispheres: isometry for the posterior vermis ( $a_{FAS} = 1.18$ ;  $p = .237$ ) and positive allometry for the inferior vermis ( $a_{FAS} = 1.51$ ;  $p < .001$ ).

The effect of FAS diagnosis as a covariate on the scaling exponent was significant for every volume and sub-volume ( $p_{DIAG} < 0.05$ ) (Figure 3).

### 3.3 | Normative analysis: Undersized volumes in FAS

In FAS, the cerebellum showed an excessively small volume relative to the whole brain one, an undersizing that did not distribute equally between cerebellar sub-volumes. The “deviation from typical scaling” ( $\sqrt{DTS}$ ) was used to quantify the degree of volumetric undersizing considering what would be expected from the law established in controls.

The cerebellar volume was 10.6% smaller than expected considering the cerebral volume (Table 2). Within the cerebellum, a gradient of severity was identified: the  $\sqrt{DTS}$  was largely negative in the anterior hemispheres ( $-12.4\%$ ,  $p < .001$ ), null in the inferior ( $1.1\%$ ,  $p = .657$ ) and positive in the posterior hemispheres ( $+2.0\%$ ,  $p = .010$ ). Thus, there was an anterior-inferior-posterior gradient of severity of volume restriction in the hemispheres. The vermis was smaller than expected ( $-11.6\%$ ;  $p < .001$ ) and this undersizing followed the same decreasing anterior-inferior-posterior gradient ( $-16.7\%$ ,  $-9.2\%$ ,  $-8.6\%$ ,  $p < .001$ ) (Table 2). The  $\sqrt{DTS}$  were significantly different two-by-two ( $p < .001$ ), except the posterior vermis with the

inferior vermis ( $p = .355$ ) and the posterior hemispheres with the inferior hemispheres ( $p = .273$ ). The average antero-infero-posterior gradient identified in subjects with FAS is shown using a colored scale on a subject with FAS's cerebellar segmentation (Figure 4C).

## 3.4 | Cerebellar-based classifiers

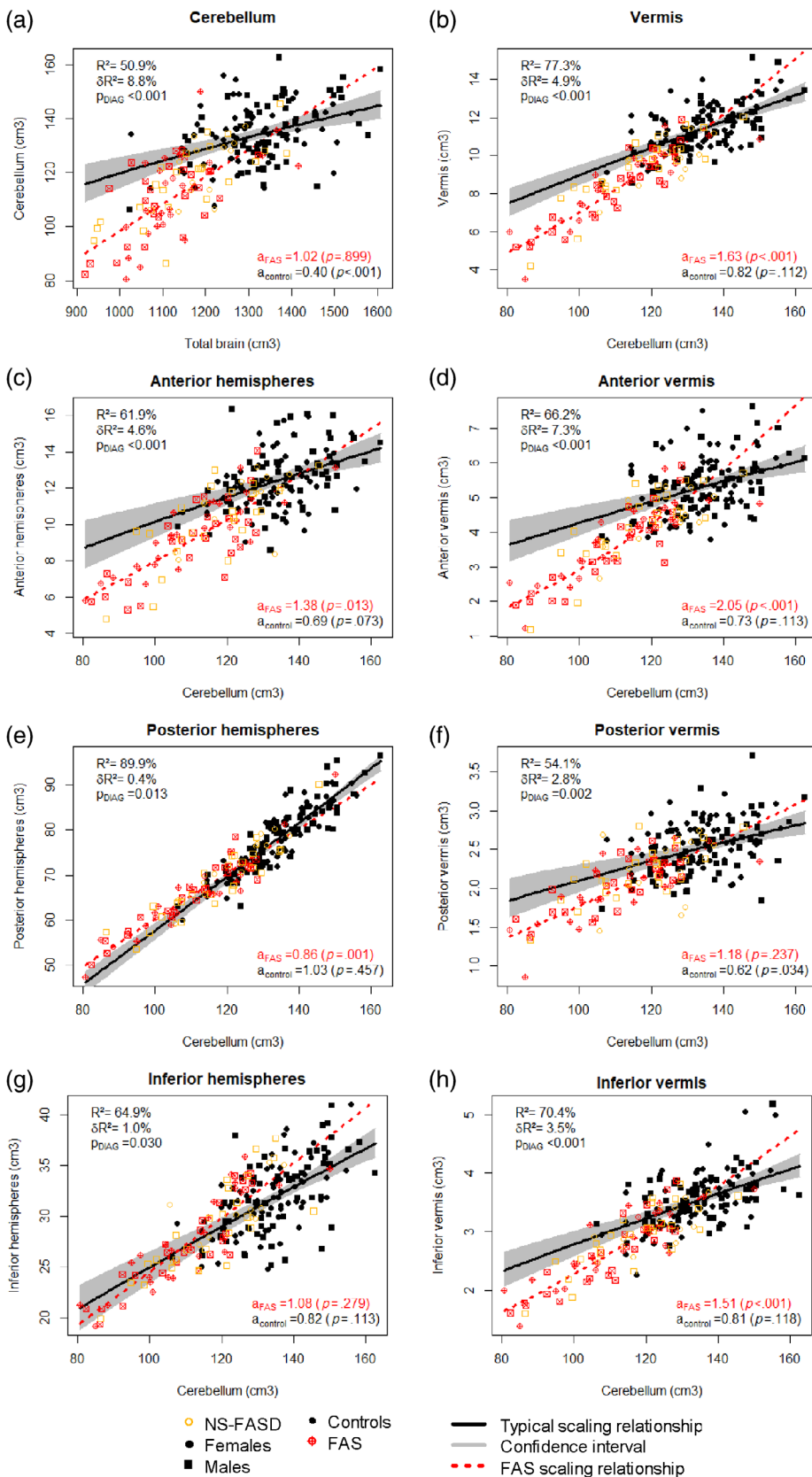
### 3.4.1 | Classification efficiency: Subjects with FAS versus typically developing controls

The logistic model including the six intracerebellar  $\sqrt{DTS}$  was significantly more efficient in explaining the diagnostic status (FAS or control) than the one based on the cerebellar  $\sqrt{DTS}$  only ( $p_{WALD} < 0.001$ , lower AIC and BIC).

The ROC curves established for the two classifiers were used to define the  $AUC_{global}$  and the classifier including the intracerebellar  $\sqrt{DTS}$  (gradient-based classifier) was found to perform better than the one based on the cerebellum  $\sqrt{DTS}$  only (cerebellum-only classifier) ( $AUC_{global} = 92\%$  vs.  $82\%$ ,  $p < .001$ ). The final threshold probability for a specificity of FAS prediction of at least  $>95\%$  was 0.7 for both classifiers used in the whole population.

The five-fold cross-validation confirmed the higher AUC for the gradient-based classifier ( $AUC_{CV} = 91\%$  vs.  $83\%$ ). The estimated accuracy was also improved by the adjunction of the intracerebellar  $\sqrt{DTS}$  ( $84\%$  vs.  $77\%$ ) (Figure 5).





**FIGURE 3** Brain scaling charts of cerebellar volumes with total brain volume (a) and cerebellar scaling charts of cerebellar sub-volumes with total cerebellum volume (B to H), in the FAS group (red) compared to the control group (black). NS-FASD represented in orange. Bottom left: Scaling coefficients  $a_{FAS}/a_{Controls}$ . Top right: Determination coefficient  $R^2$  and  $\delta R^2$ , FAS diagnosis effect ( $p_{DiAG}$ ). Confidence interval (95%) of fitted model in the control group in grey. All  $p$ -values are adjusted to achieve a  $FDR = 5\%$ .

### 3.4.2 | Application to the NS-FASD population

The two classifiers trained in the FAS and control samples with the classification threshold set for high diagnostic specificity (>95%) correctly

classified 27% and 60% of subjects with FAS respectively (Figure 6a). The use of the gradient-based classifier resulting in a 33% increase in well-classified subjects with FAS. When applied to the NS-FASD sample, the classifier including intracerebellar  $\nu$ DTS identified 13 NS-FASD (35%)

**TABLE 2** Deviation from typical scaling ( $\sqrt{\text{DTS}}$ ) of cerebellar volume and sub-volumes, in subjects with FAS, with 95% confidence interval estimated by bootstrap resampling.

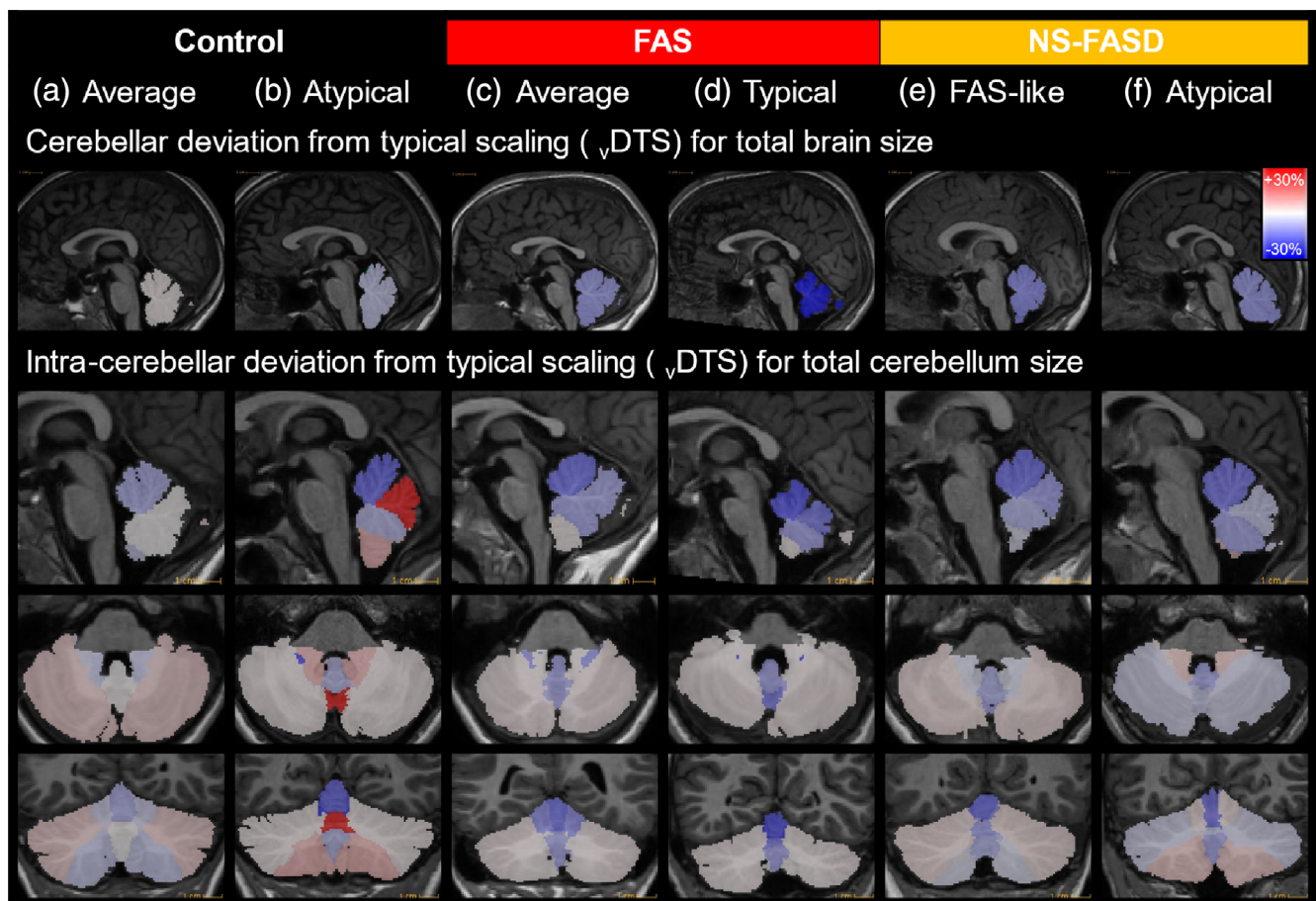
	$\sqrt{\text{DTS}}$ (%)	IC95%	<i>p</i> -value
Cerebellum	-10.6	[-13.2; -8.1]	<.001
Anterior hemispheres	-12.4	[-15.4; -9.3]	<.001
Posterior hemispheres	2.0	[1.0; 3.0]	.010
Inferior hemispheres	1.1	[-0.8; 3.0]	.657
Vermis	-11.6	[-14.1; -9.1]	<.001
Anterior vermis	-16.7	[-20.3; -13.1]	<.001
Posterior vermis	-8.6	[-11.4; -5.9]	<.001
Inferior vermis	-9.2	[-11.9; -6.4]	<.001

Note: Comparison of the FAS and controls  $\sqrt{\text{DTS}}$  by embedded bootstrap resampling. All *p*-values are adjusted to achieve a FDR = 5%, in bold *p*-values <0.05.

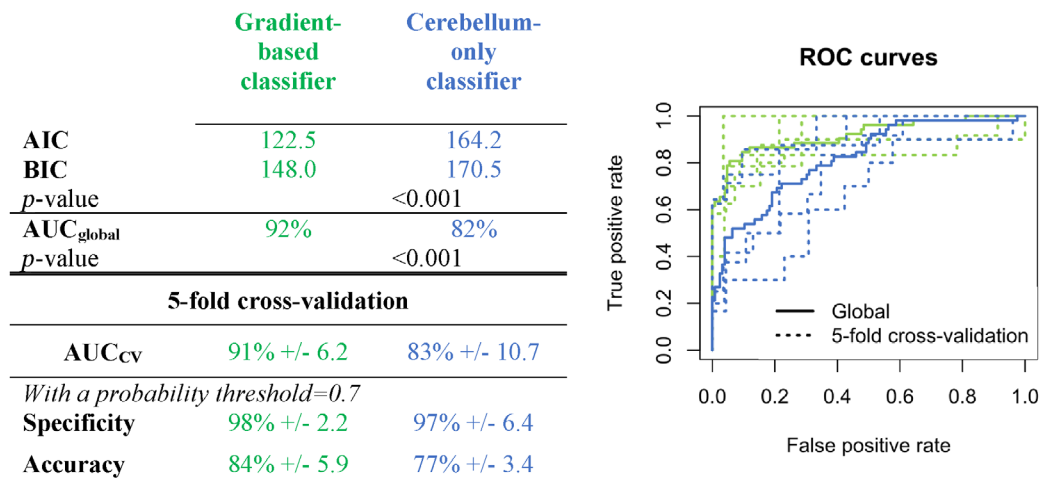
as probable FAS, compared to only 4 (11%) on the basis of total cerebellum  $\sqrt{\text{DTS}}$  alone ( $p_{\text{FISHER}} = 0.027$ ) (see subject E in Figure 4).

## 4 | DISCUSSION

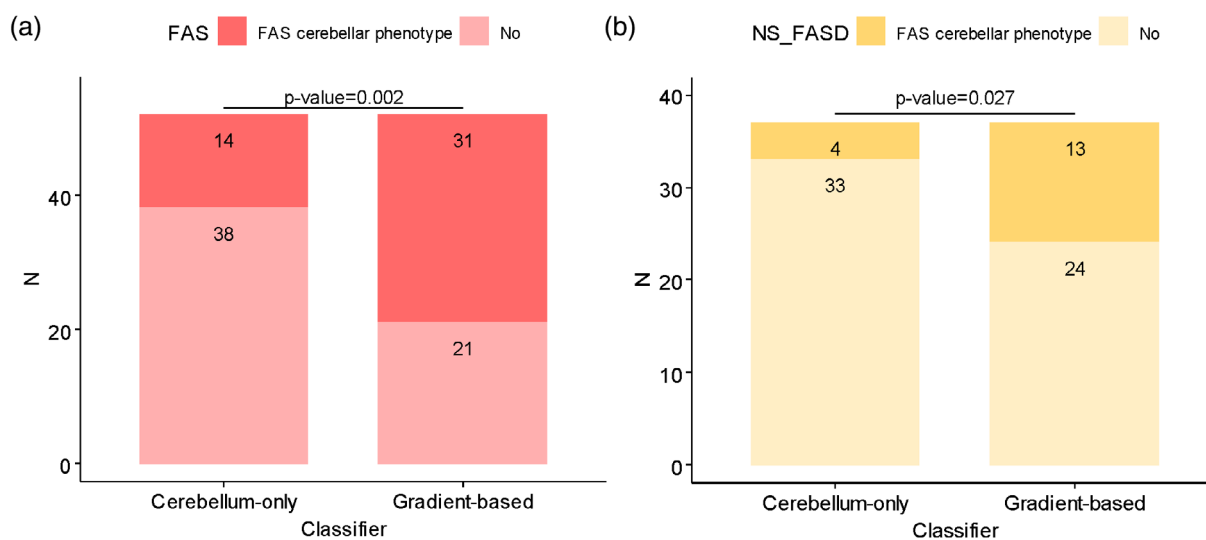
Our study is the first to compare the allometric variations of cerebellar lobar and vermian volumes in 126 typically developing controls and 89 subjects with FASD aged 6–20 years. Our results confirm that the cerebellum is particularly affected by volumetric undersizing in FAS compared to typically developing subjects and describe an intracerebellar gradient of volumetric undersizing from over-impaired vermis and anterior regions to less affected inferior and posterior ones. Used as a multidimensional neuroanatomical marker in a logistic classifier, this gradient is more efficient than the sole cerebellar volume undersizing to discriminate FAS from non-FAS healthy controls. Trained in the FAS and control population, the gradient-based classifier also



**FIGURE 4** Mapping the “deviation from typical scaling” ( $\sqrt{\text{DTS}}$ ) in selected subjects. (a) Control subject close to the norm (almost null  $\sqrt{\text{DTS}}$ ). (b) Control subject with small cerebellum but without the specific gradient of FAS. (c) Average antero-infero-posterior gradient of Subjects with FAS (Table 2). (d) Subject with FAS with obvious dysgenesis of the anterior vermis; the typical antero-infero-posterior gradient is amplified. (e) Subject with NS-FASD with the same typical antero-infero-posterior gradient (classified as FAS). (f) Subject with NS-FASD with a small vermis for the size of the cerebellum but without the specific gradient (not classified as FAS). Note the similarity of the gradient between C, D, and E.



**FIGURE 5** Comparison of FAS vs. controls classifiers using either the cerebellar vDTS only (blue) or adding the six intracerebellar vDTS (green). Left: AIC, BIC and *p*-value of the Wald test. AUC and *p*-value of comparison with the DeLong method. Right: ROC curves of each classifier in the whole sample (solid line) and in the 5-fold cross-validation samples (dotted), mean of the performance indices of the 5-fold cross-validation with the standard deviation. AIC, Akaike information criterion; AUC, area under the curve; BIC, Bayesian information criterion; DTS, deviation from typical scaling; ROC, receiver operating characteristic.



**FIGURE 6** Proportion of FAS (a) and NS-FASD (b) classified as FAS cerebellar phenotype with the cerebellum-only classifier (left, dark color) and gradient-based classifier (right, dark color), with the classification threshold of 0.7 corresponding to a 95% diagnostic specificity.

identifies more than a third of NS-FASD as having a FAS-like pattern of intracerebellar volumetric undersizing.

#### 4.1 | Presence and significance of an intracerebellar gradient of volumetric undersizing in subjects with FAS/FASD

The historical descriptions of vermis abnormalities in FASD (O'Hare et al., 2005; Sowell et al., 1996) showed a smaller area of the vermis than control, on the sagittal midsection ( $N = 9$  subjects with PAE and  $N = 21$  subjects with PAE and FAS, respectively). Cardenas

et al. (2014) strongly suggested the whole cerebellum and hemispheric volume to be undersized, taking into account brain size (intracranial volume) linearly in a small cohort of subjects ( $N = 11$  subjects with PAE, no correction for multiple comparisons). In a sizeable cohort, Astley et al. (2009) also highlighted a smaller vermian area, affecting more severely its anterior part. Despite small sample sizes or sometimes simple tools, these studies already pointed to both anterior damage of the cerebellar vermis with relative preservation of its posterior part and more severe damage in subjects with FAS than in "affected" subjects without facial features. In the present study on 89 subjects with FASD (52 FAS), we confirm that the whole cerebellum is smaller than expected in subjects with FAS (−10.6%),

considering non-linear allometric scaling. We also clearly demonstrate that this volumetric undersizing, compared to what would be expected on the basis of the controls, is not homogeneously distributed within the cerebellum between anterior, posterior, and inferior regions of the vermis and the hemispheres: the most affected regions are the anterior vermis (−16.7%) then the anterior hemispheres (−12.4%), but also the rest of the vermis (−11.6% as a whole), while the inferior and posterior hemispheres are less affected if not relatively preserved respectively (Table 2 and Figure 3). Therefore, we propose to consider the cerebellar damage as a complex neuroanatomical phenotype reflecting an intracerebellar anterior-inferior-posterior gradient of volumetric undersizing both at the vermic and hemispheric levels and characterizing the pathologically small cerebellum of subjects with FAS (Figure 4).

This gradient of cerebellar damage is probably the result of a prenatal extreme alcohol sensitivity of cerebellar cells. The reasons for the spatial heterogeneity of damage are to be sought in experimental science and animal models. A recent review of studies in murine models (Almeida et al., 2020) confirmed that all developmental stages of the cerebellum (gliogenesis, cell differentiation, synaptogenesis and myelination) could be hindered by prenatal ethanol exposure. Studies of animal models, both murine and ovine, have shown that PAE, whether occasional or repeated, results in massive loss of Purkinje cells in the vermis in utero (Idrus & Napper, 2012; Marcussen, Goodlett, Mahoney, & West, 1994; Nirgudkar, Taylor, Yanagawa, & Valenzuela, 2016; Sawant et al., 2013), and a lower volume of the vermic lobules, with differences in affected regions depending on the term at exposure (Nirgudkar et al., 2016; Sawant et al., 2013). However, these few experiments did not consider the global deficit in size, nor did they systematically investigate the whole cerebellum. They were thus unable to show a specifically vermic damage, nor did they systematically show a more anterior cerebellar damage. One hypothesis, in addition to an overall sensitivity of Purkinje cells, could be a graduated sensitivity related to differential cell maturation, such as in the posterior cerebellar hemispheres, which mature earlier and may be therefore less sensitive to cell loss (Sawant et al., 2013). In human studies, however, the timing and level of exposure at each stage of pregnancy, retrospectively and/or by statement, are necessarily biased or inaccessible, making assumptions about a possible link illusory.

Understanding the significance of cerebellar damage observed after PAE also means interrogating the functional consequences to be expected. FASD subjects have cognitive profiles that variably combine impairments in cognitive functions (intellectual efficiency, coordination, language, executive functions, etc.) as well as in emotional regulation, often associated in a complex multidimensional disorder (Lange, Rovet, Rehm, & Popova, 2017). Interestingly, these NDDs, although non-specific, could be related to the affective-cognitive cerebellar syndrome described by Schmahmann & Sherman Schmahman and Sherman (1998) and Schmahmann Schmahmann (2010) and to the now well-known cognitive role of the cerebellum (Schmahmann, 2019). A developmental form of this functional syndrome has been proposed by some authors (Brossard-Racine &

Limperopoulos, 2016). However, functional studies on the cerebellum strongly suggest that the “cognitive cerebellum” is mainly located in the posterior part of the cerebellum, and that the anterior cerebellum mostly has a motor role (Stoodley & Schmahmann, 2009) which is in apparent contradiction with our observations, but also with those of Sullivan et al. (2020). Some authors attempted to correlate anatomical damage with cognitive functioning: O'Hare et al. (2005) found a correlation between the degree of vermic displacement and verbal learning and memory functioning in subjects with PAE and Sullivan et al. (2020) found a statistical relationship between lobe VIIb volumes and intellectual quotient performance and lobe X and verbal performance. This is still quite an open field for investigations and in any case, the cerebral damage—primarily a global brain size deficit—is diffuse in subjects with FASD, so even if this size deficit is heterogenous, one could hardly expect a unicist structural-functional explanation.

## 4.2 | Towards the clinical relevance of a cerebellar signature of FAS

After ascertaining the specificity of this gradient, we were keen to see whether this identification could be of any clinically relevant use. To avoid making an obvious redundant statement, it was of course important for the neuroanatomical anomalies to be independent from the whole brain growth deficiency, considering allometric phenomena in the scaling (size effect) analysis (Bookstein et al., 2006; de Jong et al., 2017; Germanaud et al., 2012; Liu et al., 2014). Beyond theoretical concerns, we ensured that the power-law model always fitted the data at least as well as the strictly proportional model (data not shown). Overall, we observed a coefficient of determination  $R^2$  of at least 51% (Figure 4), which suggests fair adequacy of the scaling fit. It can also be noted that while we are the first to describe intracerebellar allometric laws, we confirmed the strong negative allometry of the whole cerebellum to total brain ( $a = 0.40$  in our control population) that was previously described in larger healthy cohorts ( $a = 0.54$ ) (de Jong et al., 2017; Warling et al., 2021). Lastly, we proved that the cerebellar and intracerebellar scaling laws were different in controls and in subjects with FAS for all the cerebellar and intracerebellar volumes (Figure 3).

To move from classical group comparisons towards describing individual differences with respect to a common reference model (Marquand et al., 2019; Marquand, Wolfers, Mennes, Buitelaar, & Beckmann, 2016), we established reference scaling curves similarly to long-standing growth curves in pediatric medicine (Garel et al., 2011; Rollins, Collins, & Holden, 2010). We were then able to estimate how much the observed volumes deviated from the expected ones according to the typical scaling laws (“volumetric deviation from typical scaling,”  $\sqrt{DTS}$ ) and used these distances in the scaling space as measures of the undersizing (or oversizing) of cerebellar and intracerebellar volumes in FAS. Our approach to making these deviations relevant at the individual level consisted less in a directly normative approach than in the training of a multidimensional classifier on the basis of the  $\sqrt{DTS}$ . In this way, for each subject, we obtained a diagnostic



probability of presenting with an anatomical cerebellar phenotype of FAS rather than a typical one. This classification strategy began to emerge in the field of FASD neuroanatomy and diagnosis (Little & Beaulieu, 2020) but it has mainly been used as a way to identify brain areas with a high FASD predictive weight rather than to qualify subjects. Our first step at the individual level was to show that a univariate classification based on total cerebellum  $\sqrt{\text{DTS}}$  only was already relevant for discerning FAS from controls ( $\text{AUC}_{\text{global}} = 82\%$ ), consistently with the singular isometric scaling observed in FAS. Moreover, we showed that the newly identified gradient phenotype provides valuable additional information to distinguish between FAS and typically developing subjects, leading to a significant improvement of the classifier ( $\text{AUC}_{\text{global}} = 92\%$ ,  $p < .001$ , +33% of FAS subjects correctly classified). This result argues for the use of such a multidimensional cerebellar marker that is not redundant with the whole brain size anomaly nor with the cerebellum one.

In addition to showing that the multidimensional cerebellar gradient-based classifier was more effective than the one considering only the cerebellum, we also wanted to evaluate its applicability and value at the individual level on subjects with NS-FASD. FASD imaging studies often report that subjects with non-syndromic FASD have an intermediate degree of neuroanatomical changes at the group level (comparison of means) (Archibald et al., 2001; Astley et al., 2009; Inkelis, Moore, Bischoff-Grethe, & Riley, 2020; O'Hare et al., 2005; Sullivan et al., 2020). In our cohort, subjects with NS-FASD also presented an intermediate severity of impairment (orange dots, Figure 4) in an overlapping range between subjects with FAS and typically developing controls. It was therefore logical to investigate whether what could be considered a FAS cerebellar signature, or at least a cerebellar morphology strongly suggestive of FAS, in view of the high level of specificity set for the classification (>95%), was also observed in some subjects with NS-FASD. We indeed found that the gradient-based classifier tagged 35% of the NS-FASD group as FAS-like (three times more than the 11% tagged with the cerebellum-only classifier). For all these FAS-like subjects, it seems reasonable to consider that the likelihood of a causal link with PAE is significantly higher, even in the absence of a clinical FAS (see subject E in Figure 4).

In current diagnostic procedures (Astley, 2004; Cook et al., 2016; Hoyme et al., 2016), the neuroanatomical impairments other than microcephaly that should be taken into account are not clearly listed, and in any case do not provide an argument for or against a FASD diagnosis. For subjects with NS-FASD, the link between functional impairment (NDD) and PAE remains a probabilistic diagnosis, after a thorough differential diagnosis work-up. Here, we have demonstrated that a gradient of cerebellar hemispheric and vermian damage can be used to train an efficient classifier on 178 subjects of which 30% had a positive specific FAS diagnosis (Astley, 2004). Since this classifier identified 35% subjects with NS-FASD as having a FAS-like cerebellar phenotype, we think that it advocates for a more objective and independent neuroanatomical FASD diagnosis criterion, in line with what we proposed in our previous study based on more directly available radiological-clinical measurements (Fraize et al., 2022). Were they to

be confirmed by other works, these results could provide an additional diagnostic argument and ease the way to more confident diagnoses in certain subjects presenting FASD without FAS, and therefore prompt a larger use of the full spectrum of FASD diagnosis in some medical communities where a full FAS diagnosis often remains the only option.

### 4.3 | Dealing with confounding covariates

For the sake of robustness of the fit of the reference typical scaling laws, we gathered a population of 126 controls from three different acquisition sites (sequence and imager), a third only being site- and sequence-matched to the FASD group. As expected (Takao, Hayashi, & Ohtomo, 2011), there was a site bias (see control group data, supplementary material Table A.1) that we corrected beforehand using a batch effect correcting method (Fortin et al., 2017). ComBat has proven to be a method superior to other correction methods, especially for imaging data and when using logistic regression (Fortin et al., 2018; Horng et al., 2022). However, to test the robustness of our results (potential loss of information, information bias), we performed a full *post-hoc* analysis on the raw data set (without correcting for site or sex effect), that showed very similar results in terms of gradient and reached the same conclusions on the classifier potential (supplementary material Figure A.2).

To be as rigorous and informative as possible, it would have been interesting to establish differential reference charts for females and males, since a differential sex effect is expected on volumes (Fan et al., 2010; Sullivan et al., 2020). We chose to privilege the statistical power provided by a group of 126 controls, that we would have lost by dividing the number by two. As a first precaution, we ensured that there was no difference in sex-ratio between the subgroups (Table 1). Then, we applied the same method of confounders correction as for the site. In the end, even if this choice of correction possibly erases informative sex differences, at least the result of the *post-hoc* analysis is reassuring in terms of quality of adjustment, and we trust our sex-independent results (supplementary material Figure A.2). However, a replication study on a larger sample, differentiating between females and males, would be of real interest, with the prospect of a more efficient analysis and classifier due to more finely adjusted reference scaling laws.

Lastly, choosing the scaling model and this "size-for-size" space means not including the classical normative analysis "size-for-age." Again, we first ensured that there was no difference in mean age between the groups (Table 1), and after adjusting for site and sex confounders, we found no effect of age, even with a rather biologically accurate model preserving a possible interaction with age and quadratic age (Giedd & Rapoport, 2010; Nardelli et al., 2011). This was not surprising since in the age range we studied, long after the fast growth of the first years of life and before significant adulthood shrinking, the effect of age was expected to be negligible (Inkelis et al., 2020; Jandeaux et al., 2019).

#### 4.4 | Anatomical considerations

In Schmahmann's original description of cerebellar anatomy (Schmahmann et al. 1999) the lateral boundaries of the anterior vermis were considered virtual; thus, the vermis is not individualized in many current segmentation tools (Diedrichsen et al., 2009). Yet, other neuroanatomical consensus atlases, such as the "Scalable Brain Atlas" (Bakker et al., 2015), propose a delimitation of the whole vermis lateral boundaries. Besides, the FASD literature suggests a predominantly sagittal involvement of the anterior cerebellum. That is why we proposed a combination of the currently most accurate CERES lobular segmentation (Romero et al., 2017) with a mono-atlas segmentation of the vermis optimized by the SUIT cerebellar-dedicated spatial normalization, which circumvented the lack of a completely validated tool and met our requirements to individualize the whole vermis. Given the manageable size of our cohort, systematic verification, and correction of lobar boundaries by an expert (JF) for obvious segmentation errors ensured the anatomical reliability of the studied sub-volumes, in particular when the anatomy was very abnormal (Figure 4c,d). In this context, we were not able to corroborate our results with several segmentation methods as others duly did (Biffen et al., 2020; Sullivan et al., 2020; Sullivan, Zahr, Saranathan, Pohl, & Pfefferbaum, 2019), which would have unrealistically required several combinations of tools. Among the other choices related to cerebellar anatomy, we found it unnecessary to perform a lateralized sub-volume analysis (left and right comparison). The results of the lateralized analyses of Zhou et al. (2018), confirmed by the asymmetry index study of Sullivan et al. (Sullivan et al., 2020), did not suggest cerebellar volume asymmetry in subjects exposed prenatally to alcohol. In the same spirit of a parsimonious, anatomically accurate and power-optimized study, we chose to work at the scale of the cerebellar lobes, not lobules, keeping the possibility of verification and correction of the boundaries when necessary. We also chose not to individualize the white and grey matter. In fact, the grey-white junction is sensitive to the configuration of the T1 sequence, to the maturation effect of age until late adolescence (Drakulich et al., 2021), and in any case, the cerebellar grey/white boundary is not positioned at the same place in the different cerebellar segmentations proposed for millimetric T1-weighted conventional images.

#### 4.5 | Other limitations

The retrospective design of our study led to the exclusion of a significant number of potential subjects with FASD due to incomplete data (supplementary material Figure A.1). However, it is important to emphasize that brain MRI was systematically performed on FASD-suspected subjects in the Robert-Debré reference center as part of the recommended work-up, so there was no bias was expected in terms of severity or type of subjects with FASD in our studied population. In the diagnosis procedure, we also duly considered a "relevant" threshold of PAE (Cook et al., 2016; Hoyme

et al., 2016), even if we did not have retrospective access to detailed timing and quantification. Therefore we were unable to correlate our finding with the level of PAE (Jacobson, Jacobson, Stanton, Meintjes, & Molteno, 2011; May et al., 2013). Another limitation related to the retrospective construct of this study is that MRI data were acquired in the context of clinical care, on a 1.5 T imager, which may be seen as not optimum for brain imaging. This is a grounded concern for a structure such as the cerebellum even if it should be pointed out that 60% of participants shared the same site and sequence, and that all sequences were high-quality millimetric IR-prep ultra-fast gradient echo sequences. One could also question the fact that inferior parts of the cerebellum may have suffered from a mild signal drop-off, not visible in quality-check steps, that could have tempered with our ability to reveal a group difference in this area. The last important methodological limitation, which is often found in human studies on NDD, is the choice of the control population. The comparison between subjects with FASD and typically developing controls (no alcohol, no neurological, psychiatric or learning condition) appears as a necessary step, but to ensure a real specificity of any cerebellar phenotype in FASD, it would be necessary to compare with a population of subjects with a non-alcohol related NDD and/or a population of "small brains" not related to alcohol (any type of microcephaly of paired severity). Future studies consequently should focus on separate female/male groups, use multiple segmentation tools, and compare results with various samples of subjects with NDD or small brain populations, as to the best of the authors' knowledge this has not yet been proposed in other comparable studies.

## 5 | CONCLUSION

Our study is the first to consider allometric scaling to describe the cerebellar damage in a large cohort of FASD, revealing an anterior-inferior-posterior gradient of sensitivity to PAE both in hemispheres and vermis. More informative than cerebellar undersizing alone, the intracerebellar gradient could constitute an advance towards a first neuroanatomical signature of FAS to be searched in NS-FASD to try and improve the specificity of the diagnosis.

#### ACKNOWLEDGMENTS

The authors would like to thank the volunteers, patients and families, and the French supportive association for FASD-affected families "Vivre avec le SAF."

#### FUNDING INFORMATION

This study was supported by the French National Agency for Research (ANR-19-CE17-0028-01) and the French National Institute for Public Health research (IRESP-19-ADDICTIONS-08).

#### CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to disclose.



## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

## ORCID

Justine Fraize  <https://orcid.org/0000-0001-6434-7992>

Edouard Duchesnay  <https://orcid.org/0000-0002-4073-3490>

## REFERENCES

- Almeida, L., Andreu-Fernández, V., Navarro-Tapia, E., Aras-López, R., Serra-Delgado, M., Martínez, L., García-Algar, O., & Gómez-Roig, M. D. (2020). Murine models for the study of fetal alcohol spectrum disorders: An overview. *Frontiers in Pediatrics*, 8, 359. <https://doi.org/10.3389/fped.2020.00359>
- Archibald, S. L., Fennema-Notestine, C., Gamst, A., Riley, E. P., Mattson, S. N., & Jernigan, T. L. (2001). Brain dysmorphology in individuals with severe prenatal alcohol exposure. *Developmental Medicine and Child Neurology*, 43, 148–154. <https://doi.org/10.1111/j.1469-8749.2001.tb00179.x>
- Astley, S. J. (2004). *Diagnostic guide for fetal alcohol spectrum disorders: The 4-digit diagnostic code* (3rd ed.). University of Washington.
- Astley, S. J., Aylward, E. H., Olson, H. C., Kerns, K., Brooks, A., Coggins, T. E., Davies, J., Dorn, S., Gendler, B., Jirikowic, T., Kraegel, P., Maravilla, K., & Richards, T. (2009). Magnetic resonance imaging outcomes from a comprehensive magnetic resonance study of children with fetal alcohol spectrum disorders. *Alcoholism, Clinical and Experimental Research*, 33, 1671–1689. <https://doi.org/10.1111/j.1530-0277.2009.01004.x>
- Bakker, R., Tiesinga, P., & Kötter, R. (2015). The scalable brain atlas: Instant web-based access to public brain atlases and related content. *Neuroinformatics*, 13, 353–366. <https://doi.org/10.1007/s12021-014-9258-x>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B: Methodological*, 57, 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Biffen, S. C., Warton, C. M. R., Dodge, N. C., Molteno, C. D., Jacobson, J. L., Jacobson, S. W., & Meintjes, E. M. (2020). Validity of automated FreeSurfer segmentation compared to manual tracing in detecting prenatal alcohol exposure-related subcortical and corpus callosal alterations in 9- to 11-year-old children. *NeuroImage Clinical*, 28, 102368. <https://doi.org/10.1016/j.nicl.2020.102368>
- Bookstein, F. L., Streissguth, A. P., Connor, P. D., & Sampson, P. D. (2006). Damage to the human cerebellum from prenatal alcohol exposure: The anatomy of a simple biometrical explanation. *Anatomical Record. Part B, New Anatomist*, 289, 195–209. <https://doi.org/10.1002/ar.b.20114>
- Boronat, S., Sánchez-Montañez, A., Gómez-Barros, N., Jacas, C., Martínez-Ribot, L., Vázquez, E., & del Campo, M. (2017). Correlation between morphological MRI findings and specific diagnostic categories in fetal alcohol spectrum disorders. *European Journal of Medical Genetics*, 60, 65–71. <https://doi.org/10.1016/j.ejmg.2016.09.003>
- Bouyeure, A., Germanaud, D., Bekha, D., Delattre, V., Lefèvre, J., Pinabiaux, C., Mangin, J.-F., Rivière, D., Fischer, C., Chiron, C., Hertz-Pannier, L., & Noulhiane, M. (2018). Three-dimensional probabilistic maps of mesial temporal lobe structures in children and Adolescents' brains. *Frontiers in Neuroanatomy*, 12, 98. <https://doi.org/10.3389/fnana.2018.00098>
- Brossard-Racine, M., & Limperopoulos, C. (2016). Normal cerebellar development by qualitative and quantitative MR imaging. *Neuroimaging Clinics of North America*, 26, 331–339. <https://doi.org/10.1016/j.nic.2016.03.004>
- Carass, A., Cuzzocreo, J. L., Han, S., Hernandez-Castillo, C. R., Rasser, P. E., Ganz, M., Beliveau, V., Dolz, J., Ben Ayed, I., Desrosiers, C., Thyreau, B., Romero, J. E., Coupé, P., Manjón, J. V., Fonov, V. S., Collins, D. L., Ying, S. H., Onyike, C. U., Crocetti, D., ... Prince, J. L. (2018). Comparing fully automated state-of-the-art cerebellum parcellation from magnetic resonance images. *NeuroImage*, 183, 150–172. <https://doi.org/10.1016/j.neuroimage.2018.08.003>
- Cardenas, V. A., Price, M., Infante, M. A., Moore, E. M., Mattson, S. N., Riley, E. P., & Fein, G. (2014). Automated cerebellar segmentation: Validation and application to detect smaller volumes in children prenatally exposed to alcohol. *NeuroImage Clinical*, 4, 295–301. <https://doi.org/10.1016/j.nicl.2014.01.002>
- Cook, J. L., Green, C. R., Lilley, C. M., Anderson, S. M., Baldwin, M. E., Chudley, A. E., Conry, J. L., LeBlanc, N., Loock, C. A., Lutke, J., Mallon, B. F., McFarlane, A. A., Temple, V. K., & Rosales, T. (2016). Fetal alcohol spectrum disorder: A guideline for diagnosis across the lifespan. *Canadian Medical Association Journal*, 188, 191–197. <https://doi.org/10.1503/cmaj.141593>
- de Jong, L. W., Vidal, J.-S., Forsberg, L. E., Zijdenbos, A. P., Haight, T., Alzheimer's Disease Neuroimaging Initiative, Sigurdsson, S., Gudnason, V., van Buchem, M. A., & Launer, L. J. (2017). Allometric scaling of brain regions to intra-cranial volume: An epidemiological MRI study. *Human Brain Mapping*, 38, 151–164. <https://doi.org/10.1002/hbm.23351>
- DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*, 44, 837–845. <https://doi.org/10.2307/2531595>
- Diedrichsen, J., Balsters, J. H., Flavell, J., Cussans, E., & Ramnani, N. (2009). A probabilistic MR atlas of the human cerebellum. *NeuroImage*, 46, 39–46. <https://doi.org/10.1016/j.neuroimage.2009.01.045>
- Drakulich, S., Thiffault, A.-C., Olafson, E., Parent, O., Labbe, A., Albaugh, M. D., Khundrakpam, B., Ducharme, S., Evans, A., Chakravarty, M. M., & Karama, S. (2021). Maturational trajectories of pericortical contrast in typical brain development. *NeuroImage*, 235, 117974. <https://doi.org/10.1016/j.neuroimage.2021.117974>
- Fan, L., Tang, Y., Sun, B., Gong, G., Chen, Z. J., Lin, X., Yu, T., Li, Z., Evans, A. C., & Liu, S. (2010). Sexual dimorphism and asymmetry in human cerebellum: An MRI-based morphometric study. *Brain Research*, 1353, 60–73. <https://doi.org/10.1016/j.brainres.2010.07.031>
- Fortin, J.-P., Cullen, N., Sheline, Y. I., Taylor, W. D., Aselcioglu, I., Cook, P. A., Adams, P., Cooper, C., Fava, M., McGrath, P. J., McInnis, M., Phillips, M. L., Trivedi, M. H., Weissman, M. M., & Shinohara, R. T. (2018). Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage*, 167, 104–120. <https://doi.org/10.1016/j.neuroimage.2017.11.024>
- Fortin, J.-P., Parker, D., Tunç, B., Watanabe, T., Elliott, M. A., Ruparel, K., Roalf, D. R., Satterthwaite, T. D., Gur, R. C., Gur, R. E., Schultz, R. T., Verma, R., & Shinohara, R. T. (2017). Harmonization of multi-site diffusion tensor imaging data. *NeuroImage*, 161, 149–170. <https://doi.org/10.1016/j.neuroimage.2017.08.047>
- Fraize, J., Garzón, P., Ntorkou, A., Kerdreux, E., Boespflug-Tanguy, O., Beggiano, A., Delorme, R., Hertz-Pannier, L., Elmaleh-Berges, M., & Germanaud, D. (2022). Combining neuroanatomical features to support diagnosis of fetal alcohol spectrum disorders. *Developmental Medicine and Child Neurology*, 65, 551–562. <https://doi.org/10.1111/dmcn.15411>
- Garel, C., Cont, I., Alberti, C., Josserand, E., Moutard, M. L., & Ducou le Pointe, H. (2011). Biometry of the corpus callosum in children: MR imaging reference data. *American Journal of Neuroradiology*, 32, 1436–1443. <https://doi.org/10.3174/ajnr.A2542>
- Germanaud, D., Lefèvre, J., Fischer, C., Bintner, M., Curie, A., des Portes, V., Eliez, S., Elmaleh-Bergès, M., Lamblin, D., Passemard, S., Operto, G., Schaefer, M., Verloes, A., Toro, R., Mangin, J. F., & Hertz-Pannier, L. (2014). Simplified gyral pattern in severe developmental microcephalies? New insights from allometric modeling for spatial and

- spectral analysis of gyrification. *NeuroImage*, 102(Pt 2), 317–331. <https://doi.org/10.1016/j.neuroimage.2014.07.057>
- Germanaud, D., Lefèvre, J., Toro, R., Fischer, C., Dubois, J., Hertz-Pannier, L., & Mangin, J.-F. (2012). Larger is twistier: Spectral analysis of gyrification (SPANGY) applied to adult brain size polymorphism. *NeuroImage*, 63, 1257–1272. <https://doi.org/10.1016/j.neuroimage.2012.07.053>
- Giedd, J. N., & Rapoport, J. L. (2010). Structural MRI of pediatric brain development: What have we learned and where are we going? *Neuron*, 67, 728–734. <https://doi.org/10.1016/j.neuron.2010.08.040>
- Hong, H., Singh, A., Yousefi, B., Cohen, E. A., Haghighi, B., Katz, S., Noël, P. B., Shinohara, R. T., & Kontos, D. (2022). Generalized ComBat harmonization methods for radiomic features with multi-modal distributions and multiple batch effects. *Scientific Reports*, 12, 4493. <https://doi.org/10.1038/s41598-022-08412-9>
- Hoyme, H. E., Kalberg, W. O., Elliott, A. J., Blankenship, J., Buckley, D., Marais, A.-S., Manning, M. A., Robinson, L. K., Adam, M. P., Abdul-Rahman, O., Jewett, T., Coles, C. D., Chambers, C., Jones, K. L., Adnams, C. M., Shah, P. E., Riley, E. P., Charness, M. E., Warren, K. R., & May, P. A. (2016). Updated clinical guidelines for diagnosing fetal alcohol spectrum disorders. *Pediatrics*, 138, e20154256. <https://doi.org/10.1542/peds.2015-4256>
- Idrus, N. M., & Napper, R. M. A. (2012). Acute and long-term Purkinje cell loss following a single ethanol binge during the early third trimester equivalent in the rat. *Alcoholism, Clinical and Experimental Research*, 36, 1365–1373. <https://doi.org/10.1111/j.1530-0277.2012.01743.x>
- Inkelis, S. M., Moore, E. M., Bischoff-Grethe, A., & Riley, E. P. (2020). Neurodevelopment in adolescents and adults with fetal alcohol spectrum disorders (FASD): A magnetic resonance region of interest analysis. *Brain Research*, 1732, 146654. <https://doi.org/10.1016/j.brainres.2020.146654>
- Jacobson, S. W., Jacobson, J. L., Stanton, M. E., Meintjes, E. M., & Molteno, C. D. (2011). Biobehavioral markers of adverse effect in fetal alcohol spectrum disorders. *Neuropsychology Review*, 21, 148–166. <https://doi.org/10.1007/s11065-011-9169-7>
- Jandeaux, C., Kuchcinski, G., Ternynck, C., Riquet, A., Leclerc, X., Pruvo, J.-P., & Soto-Ares, G. (2019). Biometry of the cerebellar vermis and brain stem in children: MR imaging reference data from measurements in 718 children. *AJNR. American Journal of Neuroradiology*, 40, 1835–1841. <https://doi.org/10.3174/ajnr.A6257>
- Johnson, W. E., Li, C., & Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*, 8, 118–127. <https://doi.org/10.1093/biostatistics/kxj037>
- Lange, S., Rovet, J., Rehm, J., & Popova, S. (2017). Neurodevelopmental profile of fetal alcohol Spectrum disorder: A systematic review. *BMC Psychology*, 5, 22. <https://doi.org/10.1186/s40359-017-0191-2>
- Little, G., & Beaulieu, C. (2020). Multivariate models of brain volume for identification of children and adolescents with fetal alcohol spectrum disorder. *Human Brain Mapping*, 41, 1181–1194. <https://doi.org/10.1002/hbm.24867>
- Liu, D., Johnson, H. J., Long, J. D., Magnotta, V. A., & Paulsen, J. S. (2014). The power-proportion method for intracranial volume correction in volumetric imaging analysis. *Frontiers in Neuroscience*, 8, 356. <https://doi.org/10.3389/fnins.2014.00356>
- Manjón, J. V., & Coupé, P. (2016). volBrain: An online MRI brain volumetry system. *Frontiers in Neuroinformatics*, 10, 30. <https://doi.org/10.3389/fninf.2016.00030>
- Marcussen, B. L., Goodlett, C. R., Mahoney, J. C., & West, J. R. (1994). Developing rat Purkinje cells are more vulnerable to alcohol-induced depletion during differentiation than during neurogenesis. *Alcohol*, 11, 147–156.
- Marquand, A. F., Kia, S. M., Zabihi, M., Wolfers, T., Buitelaar, J. K., & Beckmann, C. F. (2019). Conceptualizing mental disorders as deviations from normative functioning. *Molecular Psychiatry*, 24, 1415–1424. <https://doi.org/10.1038/s41380-019-0441-1>
- Marquand, A. F., Wolfers, T., Mennes, M., Buitelaar, J., & Beckmann, C. F. (2016). Beyond lumping and splitting: A review of computational approaches for stratifying psychiatric disorders. *Biological Psychiatry Cognitive Neuroscience Neuroimaging*, 1, 433–447. <https://doi.org/10.1016/j.bpsc.2016.04.002>
- May, P. A., Blankenship, J., Marais, A.-S., Gossage, J. P., Kalberg, W. O., Joubert, B., Cloete, M., Barnard, R., de Vries, M., Hasken, J., Robinson, L. K., Adnams, C. M., Buckley, D., Manning, M., Parry, C. D. H., Hoyme, H. E., Tabachnick, B., & Seedat, S. (2013). Maternal alcohol consumption producing fetal alcohol spectrum disorders (FASD): Quantity, frequency, and timing of drinking. *Drug and Alcohol Dependence*, 133, 502–512. <https://doi.org/10.1016/j.drugalcdep.2013.07.013>
- Mosteller, F. (1968). *Data analysis, including statistics*. Handb. Soc.
- Nardelli, A., Lebel, C., Rasmussen, C., Andrew, G., & Beaulieu, C. (2011). Extensive deep gray matter volume reductions in children and adolescents with fetal alcohol Spectrum disorders: Reduced deep gray matter volume in FASD. *Alcohol: Clinical & Experimental Research*, 35, 1404–1417. <https://doi.org/10.1111/j.1530-0277.2011.01476.x>
- Nirgudkar, P., Taylor, D. H., Yanagawa, Y., & Valenzuela, C. F. (2016). Ethanol exposure during development reduces GABAergic/Glycinergic neuron numbers and lobule volumes in the mouse cerebellar vermis. *Neuroscience Letters*, 632, 86–91. <https://doi.org/10.1016/j.neulet.2016.08.039>
- O'Hare, E. D., Kan, E., Yoshii, J., Mattson, S. N., Riley, E. P., Thompson, P. M., Toga, A. W., & Sowell, E. R. (2005). Mapping cerebellar vermal morphology and cognitive correlates in prenatal alcohol exposure. *Neuroreport*, 16, 1285–1290. <https://doi.org/10.1097/01.wnr.0000176515.11723.a2>
- Rajaprakash, M., Chakravarty, M. M., Lerch, J. P., & Rovet, J. (2014). Cortical morphology in children with alcohol-related neurodevelopmental disorder. *Brain and Behavior: A Cognitive Neuroscience Perspective*, 4, 41–50. <https://doi.org/10.1002/brb3.191>
- Roebuck, T. M., Mattson, S. N., & Riley, E. P. (1998). A review of the neuroanatomical findings in children with fetal alcohol syndrome or prenatal exposure to alcohol. *Alcoholism, Clinical and Experimental Research*, 22, 339–344. <https://doi.org/10.1111/j.1530-0277.1998.tb03658.x>
- Rollins, J., Collins, J. S., & Holden, K. (2010). United States head circumference growth reference charts: Birth to 21 years. *The Journal of Pediatrics*, 156, 907–913.e2. <https://doi.org/10.1016/j.jpeds.2010.01.009>
- Romero, J. E., Coupé, P., Giraud, R., Ta, V.-T., Fonov, V., Park, M. T. M., Chakravarty, M. M., Voineskos, A. N., & Manjón, J. V. (2017). CERES: A new cerebellum lobule segmentation method. *NeuroImage*, 147, 916–924. <https://doi.org/10.1016/j.neuroimage.2016.11.003>
- Sawant, O. B., Lunde, E. R., Washburn, S. E., Chen, W.-J. A., Goodlett, C. R., & Cudd, T. A. (2013). Different patterns of regional Purkinje cell loss in the cerebellar vermis as a function of the timing of prenatal ethanol exposure in an ovine model. *Neurotoxicology and Teratology*, 35, 7–13. <https://doi.org/10.1016/j.ntt.2012.11.001>
- Schmahman, J. D., & Sherman, J. C. (1998). 1998. The cerebellar cognitive affective syndrome. *Brain: A Journal of Neurology*, 121, 561–579.
- Schmahmann, J. D. (2010). The role of the cerebellum in cognition and emotion: Personal reflections since 1982 on the dysmetria of thought hypothesis, and its historical evolution from theory to therapy. *Neuropsychology Review*, 20, 236–260. <https://doi.org/10.1007/s11065-010-9142-x>
- Schmahmann, J. D. (2019). The cerebellum and cognition. *Neuroscience Letters*, 688, 62–75. <https://doi.org/10.1016/j.neulet.2018.07.005>
- Schmahmann, J. D., Doyon, J., McDonald, D., Holmes, C., Lavoie, K., Hurwitz, A. S., Kabani, N., Toga, A., Evans, A., & Petrides, M. (1999). Three-dimensional MRI atlas of the human cerebellum in proportional stereotaxic space. *NeuroImage*, 10, 233–260. <https://doi.org/10.1006/nimg.1999.0459>
- Sowell, E. R., Jernigan, T. L., Mattson, S. N., Riley, E. P., Sobel, D. F., & Jones, K. L. (1996). Abnormal development of the cerebellar vermis in

- children prenatally exposed to alcohol: Size reduction in lobules I-V. *Alcoholism, Clinical and Experimental Research*, 20, 31–34. <https://doi.org/10.1111/j.1530-0277.1996.tb01039.x>
- Stoodley, C. J., & Schmahmann, J. D. (2009). Functional topography in the human cerebellum: A meta-analysis of neuroimaging studies. *NeuroImage*, 44, 489–501. <https://doi.org/10.1016/j.neuroimage.2008.08.039>
- Sullivan, E. V., Moore, E. M., Lane, B., Pohl, K. M., Riley, E. P., & Pfefferbaum, A. (2020). Graded cerebellar lobular volume deficits in adolescents and young adults with fetal alcohol Spectrum disorders (FASD). *Cerebral Cortex*, 30, 4729–4746. <https://doi.org/10.1093/cercor/bhaa020>
- Sullivan, E. V., Zahr, N. M., Saranathan, M., Pohl, K. M., & Pfefferbaum, A. (2019). Convergence of three parcellation approaches demonstrating cerebellar lobule volume deficits in alcohol use disorder. *NeuroImage Clinical*, 24, 101974. <https://doi.org/10.1016/j.nicl.2019.101974>
- Takao, H., Hayashi, N., & Ohtomo, K. (2011). Effect of scanner in longitudinal studies of brain volume changes. *Journal of Magnetic Resonance Imaging*, 34, 438–444. <https://doi.org/10.1002/jmri.22636>
- Toro, R., Chupin, M., Garnero, L., Leonard, G., Perron, M., Pike, B., Pitiot, A., Richer, L., Veillette, S., Pausova, Z., & Paus, T. (2009). Brain volumes and Val66Met polymorphism of the BDNF gene: Local or global effects? *Brain Structure and Function*, 213, 501–509. <https://doi.org/10.1007/s00429-009-0203-y>
- Treit, S., Jeffery, D., Beaulieu, C., & Emery, D. (2020). Radiological findings on structural magnetic resonance imaging in fetal alcohol Spectrum disorders and healthy controls. *Alcoholism, Clinical and Experimental Research*, 44, 455–462. <https://doi.org/10.1111/acer.14263>
- Treit, S., Zhou, D., Lebel, C., Rasmussen, C., Andrew, G., & Beaulieu, C. (2014). Longitudinal MRI reveals impaired cortical thinning in children and adolescents prenatally exposed to alcohol. *Human Brain Mapping*, 35, 4892–4903. <https://doi.org/10.1002/hbm.22520>
- Warling, A., McDermott, C. L., Liu, S., Seidlitz, J., Rodrigue, A. L., Nadig, A., Gur, R. C., Gur, R. E., Roalf, D., Moore, T. M., Glahn, D., Satterthwaite, T. D., Bullmore, E. T., & Raznahan, A. (2021). Regional white matter scaling in the human brain. *The Journal of Neuroscience*, 41, 7015–7028. <https://doi.org/10.1523/jneurosci.1193-21.2021>
- Zhou, D., Rasmussen, C., Pei, J., Andrew, G., Reynolds, J. N., & Beaulieu, C. (2018). Preserved cortical asymmetry despite thinner cortex in children and adolescents with prenatal alcohol exposure and associated conditions. *Human Brain Mapping*, 39, 72–88. <https://doi.org/10.1002/hbm.23818>

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Fraize, J., Fischer, C., Elmaleh-Bergès, M., Kerdreux, E., Beggiano, A., Ntorkou, A., Duchesnay, E., Bekha, D., Boespflug-Tanguy, O., Delorme, R., Hertz-Pannier, L., & Germanaud, D. (2023). Enhancing fetal alcohol spectrum disorders diagnosis with a classifier based on the intracerebellar gradient of volumetric undersizing. *Human Brain Mapping*, 1–16. <https://doi.org/10.1002/hbm.26348>